*Biomacromolecules* – version accepted for publication, February 12th 2014

# How cellulose stretches: synergism between covalent and hydrogen bonding

†Clemens M. Altaner, ‡Lynne H. Thomas, §Anwesha N. Fernandes & ‖Michael C. Jarvis

†*New Zealand School of Forestry, University of Canterbury, Christchurch 4180, New Zealand.*

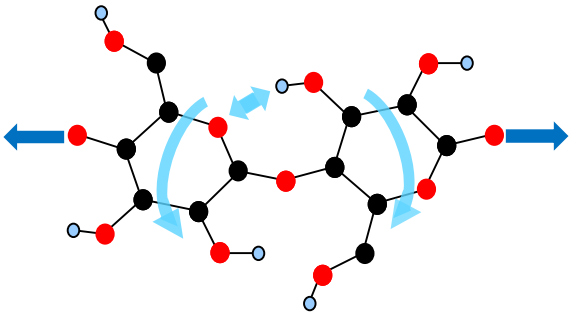
‡*Department of Chemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK.*

§*School of Physics and Astronomy, The University of Nottingham, University Park, Nottingham NG7 2RD, UK.*

‖*School of Chemistry, Glasgow University, Glasgow G12 8QQ, UK.*

**Keywords:** cellulose, elasticity; force; deformation; molecular leverage; vibrational spectroscopy; diffraction

ABSTRACT: Cellulose is the most familiar and most abundant strong biopolymer, but the reasons for its outstanding mechanical performance are not well understood. Each glucose unit in a cellulose chain is joined to the next by a covalent C-O-C linkage flanked by two hydrogen bonds. This geometry suggests some form of co-operativity between covalent and hydrogen bonding. Using infrared spectroscopy and X-ray diffraction it is shown that mechanical tension straightens out the zig-zag conformation of the cellulose chain, with each glucose unit pivoting around a fulcrum at either end. Straightening the chain leads to a small increase in its length, and is resisted by one of the flanking hydrogen bonds. This constitutes a simple form of molecular leverage with the covalent structure providing the fulcrum, and gives the hydrogen bond an unexpectedly amplified effect on the tensile stiffness of the chain. The principle of molecular leverage is directly applicable to certain other carbohydrate polymers including the animal polysaccharide chitin. Related but more complex effects are possible in some proteins and nucleic acids. The stiffening of cellulose by this mechanism is, however, in complete contrast to the way in which hydrogen bonding provides toughness combined with extensibility in protein materials like spider silk.



## INTRODUCTION

Cellulose has a simple primary structure, a linear chain of β-glucose units joined covalently by (1,4') glycosidic (C-O-C) links (Figure 1). Cellulose chains are packed into partially crystalline fibres called microfibrils, typically about 3 nm in diameter[1]. Within a microfibril the chains are arranged in sheets, with hydrogen bonding between chains and between monomers in each chain[2,3] (Figure 1). The two crystalline allomorphs cellulose Iα and Iβ are exceptionally stiff and strong, outperforming steel weight for weight[4,5] and inviting comparison with carbon nanotubes[6].

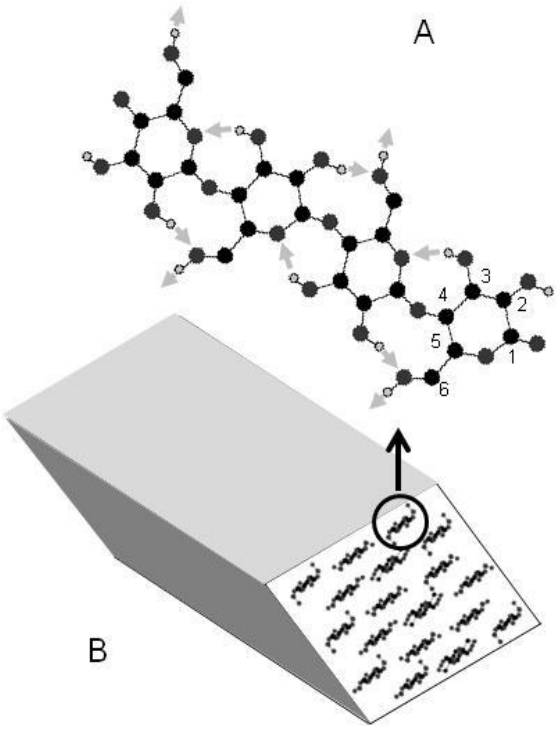



**Figure 1.** Structure of a single cellulose chain (A) showing the carbon numbering system around the glucose ring. (B) Arrangement of chains in a cellulose Iβ microfibril. Hydrogen bonds are shown as pale grey arrows.

Cellulosic materials like wood can stretch in two ways. Irreversible, time-dependent slippage can occur between the cellulose microfibrils, which reorient into line with the applied force[7]. When the force aligns with the cellulose orientation, the microfibrils themselves stretch reversibly[8]. We explored this second, elastic, stretching mechanism.

A number of modelling studies have approximately reproduced the measured elastic modulus of cellulose Iβ, ~140 GPa[9-12]. If intramolecular hydrogen bonding is eliminated from the models, the predicted tensile modulus of the cellulose decreases by up to half[9,10,12]. This prediction is unexpected because the intramolecular hydrogen bonds in cellulose are only about 10% as stiff as the covalent glycosidic linkage[11], suggesting some form of synergism between covalent and hydrogen bonding.

Experimentally, the load-bearing ability of hydrogen bonds can be investigated by vibrational spectroscopy[10]. When a hydrogen bond is stretched, the covalent O-H bond of the donor hydroxyl group is strengthened and its Fourier-transform infrared (FTIR) stretching frequency increases[11]. FTIR spectroscopy under tension, therefore, is potentially a powerful way to determine which hydrogen bonds are load-bearing and to what extent[13]. It is necessary first to be sure of the assignment of individual OH stretching bands in the FTIR spectrum to hydroxyl groups in the cellulose structure. Here we report tensile FTIR experiments supporting a proposed mechanism for the deformation of cellulose chains, that was tested by additionally carrying out wide-angle X-ray diffraction under tension.

## EXPERIMENTAL SECTION

**Materials.** Mature earlywood from Sitka spruce (*Picea sitchensis* (Bong) Carr) was prepared as described[1]. Some of the wood used in the X-ray diffraction experiments

was of Canadian origin and is thought to have been harvested for aircraft construction during the 1940s. Its microfibril angle was exceptionally low (<5°) giving tensile moduli of approximately 20 GPa at a dry density of 500-550 Kg/m$^3$. In the rest of the experiments mature wood with a microfibril angle of 6-8° was selected from the outer annual rings of Sitka spruce trees harvested in 2004 at Kershope Forest, UK[1]. For FTIR spectroscopy, longitudinal-tangential sections 20 μm in thickness were prepared wet on a sledge microtome. For X-ray diffraction, uniform longitudinal-tangential sections approximately 0.5 mm thick were prepared by hand using a straight edge and a razor blade. Particular care was necessary to align the sample axis with the longitudinal axis of the wood cells, to maintain maximum breaking strength of the samples.

**Partial internal deuteration.** Partial subsititution of cellulose hydroxyl groups with deuterium was obtained by incubating 20 μm thick longitudinal-tangential sections for 16 h in 100 mM KOH in $D_2O$ at 20°C. The KOH solution was neutralised to pH 6 with glacial acetic acid before draining and washing extensively with $H_2O$ to re-convert accessible cellulosic and non-cellulosic hydroxyl groups. The sections were dried for at least 2 days pressed between sheets of filter paper. Partial internal deuteration[14] gave O-D stretching bands that were much less subject to interference from non-crystalline domains than the O-H stretching bands remaining after vapour-phase deuteration[15]. Their lower intensity gave improved signal: noise ratio and freedom from saturation problems that hindered quantification in the intensely absorbing O-H stretching region. There was no evidence that the deuterium exchange treatment disturbed the hydrogen-bond geometry of the crystalline cellulose. Much more severe internal deuteration has been used to determine cellulose structure without altering the structure in any way[14,16].

**FTIR microscopy.** FTIR spectroscopy was carried out in transmission mode using a Thermo Nicolet Nexus Spectrometer equipped with a Nicolet Continuum microscope attachment, a liquid-nitrogen-cooled MCT detector and a wire grid polariser[1,15]. Aperture size was 100 μm in each dimension to maximise signal and minimise distortion of the spectra by scattering effects. The scanning parameters used were: resolution, 4 cm$^{-1}$; number of scans, 128. For vapour-phase deuteration experiments[1,15,17] the sample was enclosed in a through-flow cell with upper and lower $BaF_2$ windows. A stream of nitrogen, pre-dried over molecular sieve, was passed through either a drying tube filled with supported phosphorus pentoxide (Sicapent, Aldrich) or a bubbling tube filled with $D_2O$. The nitrogen line was arranged to allow switching between the drying and deuteration modes without exposure to the external atmosphere[15].

Samples were stretched progressively to breaking point on the FTIR microscope stage in a sliding rig driven directly by a M2 machine screw with 0.4 mm thread pitch. The sections, 40 mm (L) x 1 mm (T) x 20 (R) μm, were attached directly to the fixed and sliding aluminium alloy components of the rig using cyanoacrylate adhesive heat-cured for 5 min at 100ºC, taking particular care to achieve accurate axial alignment so that the stress distribution was uniform across the width of the sample. FTIR spectra were recorded as close as possible to the fixed end of the sample, so that throughout each experiment the spectra were being recorded in the same place. The exact position was limited by diffusion of the cyanoacrylate adhesive for a short distance beyond the attachment point at the fixed end of the sample. Any cyanoacrylate was readily visible in the FTIR spectra. Samples with partial internal deuteration

were examined under ambient conditions, at approximately 50% r.h. In this case the intensity of the O-H stretching bands was used to monitor the hydration status of the samples, which remained constant (± 2%) within each experiment. For vapour-phase deuteration[15] the fixed and sliding parts of the stretching rig were encased in a purpose-built cell with $BaF_2$ windows built into the top and bottom.

**Assigning O-H stretching bands in the FTIR spectra.** There is disagreement in the literature concerning these assignments, which are based on principles derived from a small number of publications dating from before the structural complexity of native cellulose was well recognised. Firstly, it was assumed that hydroxyl groups with intermolecular and intramolecular hydrogen bonds can be distinguished by the transverse and longitudinal polarisation, respectively, of their hydroxyl stretching bands[18-20]. Secondly, there has been an assumption that in native cellulose, intermolecular hydrogen bonds are shorter and therefore have lower donor O-H stretching frequencies than intramolecular hydrogen bonds[21-23]. This assumption can be traced to two independent studies[11,24] based on hydrogen-bond lengths[25] from the Gardner and Blackwell structure[26], which are indeed shorter for the intermolecular than the intramolecular hydrogen bonds. However no such pattern is evident in the currently accepted structures[16,27] for cellulose Iα and Iβ, now known to be distinct. Further, recent DFT modelling studies[28] have shown that coupling is quite extensive between O-H stretching modes within each chain. Coupling leads to averaging of the polarisation vectors of the coupled modes and brings all polarisation closer to neutral. In their current form the DFT modelling studies[28] do not accurately predict absolute frequencies. We therefore used the order of the DFT-predicted bands in the spectrum, together with observed polarisation data, to assign the FTIR spectra. In particular this gave a clear assignment of the 2441 $cm^{-1}$ band in the spectrum of internally deuterated cellulose to a coupled stretching vibration of O2-D and O6-D in the A network of cellulose Iβ, with its longitudinally polarised component attributable mainly to O2-D stretching. Making allowance for the OD/OH frequency ratio, this assignment agrees with published assignments based on polarisation[19], but not those based on hydrogen-bond length[21-23].

**Measuring FTIR frequency shifts.** Spectra were baseline-corrected using a segmented linear baseline joining the following frequencies: 843, 1550, 1818, 2289, 2635, 3303 and 3764 $cm^{-1}$. The extent of cellulose stretching was estimated from the frequency shift of the 1162 $cm^{-1}$ band in the longitudinally polarised spectra, calculated as follows. The longitudinally polarised spectra from a single stretching experiment were normalised on the intensity at 1162 $cm^{-1}$ and averaged. Each normalised spectrum was then matched against the averaged spectrum shifted in frequency by a variable amount $\delta\nu$. Least-squares minimisation was then used to optimise $\delta\nu$. The negative frequency shift at 1162 $cm^{-1}$ was in general linear with macroscopic strain, and experiments in which this relationship was significantly non-linear were discarded.

Frequency shifts in the O-D stretching region (2400-2600 $cm^{-1}$) were determined in two independent ways. First, the frequency shift at the maximum of the O-D stretching region, 2490 $cm^{-1}$, was measured as described above for the 1162 $cm^{-1}$ band. A slightly modified version of the difference integral method[29] was then used to calculate the local bandshift at each frequency across the O-D stretching region. Instead of integrating differences numerically from the baseline point at one end of the O-D stretching region to the other, as previously described[29], the integration was done outwards in each direction from the maximum at 2490 $cm^{-1}$,

using the frequency shift already calculated at 2490 $cm^{-1}$ as the integration constant. This change in procedure minimised random errors across the whole frequency range studied, confining them to frequencies close to the 2490 $cm^{-1}$ maximum. That is why there is a short gap in the spectral plot of bandshifts (Figure 2) between 2480 and 2500 $cm^{-1}$.

The second approach was to deconvolute the O-D stretching region of the spectrum into individual bands[17]. In principle the number of bands present should be very large, with six crystallographically distinct hydroxyl groups in the unit cell of each allomorph, cellulose Iα and Iβ; two different hydrogen-bonding networks for each allomorph; and multiple coupled vibrational modes differing in phase combinations. In practice the O-D stretching region of the samples after partial internal deuteration was consistent with a preponderance of cellulose Iβ in the A network form, and DFT modelling[28] indicates that coupled modes are grouped in frequency with fewer groups than the number of crystallographically distinct hydroxyl groups. Making use of the transversely and longitudinally polarised spectra together as described[17], it was possible to obtain a robust separation into five Gaussian bands corresponding approximately to those identified for cellulose Iβ[17], plus two broader bands at 2412 $cm^{-1}$ and 2538 $cm^{-1}$ that probably included both diffuse intensity (the high- and low-frequency tails) and contributions from cellulose Iα and the B hydrogen-bond network of cellulose Iβ. Their presence implied that some additional intensity from these other forms of cellulose probably also underlies the rest of the O-D stretching region and complicates the separation into the five bands shown, but modelling further minor bands would have introduced too many adjustable parameters to permit robust fits. Band fitting was done by least-squares minimisation using the *Solver* function in Microsoft Excel. The best-fit band frequencies, widths and intensities were first calculated for the averaged, normalised spectra from each experiment, and the shifted frequencies for all the bands were then fitted for each normalised spectrum while holding the widths and intensities constant. The broad bands at 2412 $cm^{-1}$ and 2538 $cm^{-1}$ were excluded from the statistical analysis. For the remaining deconvoluted bands, bandshifts from five experiments were averaged and significant differences were identified by one-way ANOVA.

**X-ray Diffraction.** X-ray diffraction patterns were obtained at ambient temperature and humidity (50% r.h. approximately) using a Rigaku R-axis/RAPID image plate diffractometer. A Mo Kα ($\lambda$ = 0.07071 nm) source was used, with the beam collimated to a diameter of 0.5 mm[30]. Spruce samples about 0.5 mm thick in the direction parallel to the beam and 2 mm wide were bonded at each end between 0.5 mm aluminium alloy tags using metal-filled epoxy resin, heat-cured for 10 min at 105ºC. The sample was attached at each end to a stretching rig on the goniometer head of the diffractometer, by titanium pins fitted through 3mm holes in the aluminium tags. The sample was stretched by a M2 machine screw driving a lever arm giving 6:1 leverage. The diffraction patterns were collected from a point close to the fixed end of the sample, normally in perpendicular transmission mode. In principle tilting experiments are preferred when measuring axial reflections from crystalline fibres. However the 004 axial reflections from wood cellulose was readily observable without tilting, and in tilting experiments where this reflection could be measured only at one end of the meridian its position could not be determined with quite as much accuracy due to small deviations in the centring of the diffraction pattern during each stretching experiment. Both tilting and non-tilting modes were therefore used to collect the unit cell dimensions presented

here, but only tilting mode was used to measure any changes in the radial width of the 004 reflection that might indicate redistribution of stress between microfibrils when the sample was stretched. This was necessary because when the sample is not tilted, a disproportionate fraction of the 004 intensity is likely to be derived from slightly misoriented microfibrils.

The Rigaku CrystalClear 1.4.0 and AreaMax 1.1.5 software packages (Rigaku Inc., 9009 New Trails Dr., The Woodlands, Texas 77381, USA) were used to collect and process images. Each diffraction pattern, corrected for detector geometry but without background subtraction, was extracted in the form of 180 radial profiles each integrated over 2º of the azimuthal angle $\chi$[30]. To construct difference diffraction patterns it was necessary first to equalise the rotation and centring of the diffraction patterns by least-squares minimisation of differences in the intensity of the 200 reflection between quadrants, using the *Solver* function in Microsoft Excel.

The positions of the 200, 1-10 and 110 reflections were then determined from the radial intensity profiles averaged over 20º in azimuth. This wide azimuthal distribution was necessary to capture the whole azimuthal range covered by each reflection, because some redistribution of intensity from the wings into the centre of the reflection occurred on stretching as microfibrils realigned towards the strain axis.

Any remaining discrepancies in centring were corrected by averaging the distance from the centre ($2\theta$) for each pair of opposite reflections. External calibration of $2\theta$ was checked with $LaB_6$[30].

Calculation of the axial (*c*) dimension of the unit cell was based on $2\theta$ for the 004 reflections, each fitted first in the azimuthal and then in the radial direction as described[30] using a Gaussian radial profile over a linear local baseline.

The positions of the equatorial 1-10, 110, and 200 reflections were each fitted with a linear baseline and an asymmetric function of the form[15]:

$$I = kI_o(1+f(2\theta)) \exp(-0.5((2\theta-2\theta_o)/\sigma)^{1.8}/\sigma$$

Where $k$ is a scaling constant, $I_o$ is the maximum intensity located at $2\theta = 2\theta_o$ $\sigma$ is a constant describing the radial width of the reflection and $f(2\theta) = 0.3(2\theta-2\theta_o)^2$ when $2\theta < 2\theta_o$ but zero when $2\theta > 2\theta_o$. The use of an asymmetric modelled profile allowed closer fits to the data than any symmetric function. The *a* dimension of the unit cell was calculated directly from the fitted position of the 200 reflection. The *b* dimension and the monoclinic angle $\gamma$ were then calculated simultaneously from the fitted positions of the 1-10 and 110 reflections, by a numerical adaptation of a method described previously[31]. For each experiment the mean values of *a*, *b*, *c* and $\gamma$ were calculated, and hence the percent deviation from these mean values at each strain level was obtained. This allowed the data from six experiments to be pooled ($n$ = 27), and *a*, *b*, and $\gamma$ were subjected separately to regression against *c* for the combined data sets.

**Tensile testing.** The samples were the same as those used for X-ray diffraction, with aluminium alloy tags at the ends. Load-deformation and stress-relaxation curves were collected on a Tinius Olsen H1KS testing instrument, correcting displacement of the crosshead for instrumental deflection, and were converted to stress and engineering strain using sample dimensions measured to ± 5 μm with a digital micrometer.

## RESULTS

FTIR spectra were collected under tension from thin foils of Sitka spruce wood, less than one cell thick. The wood used was selected for microfibril orientation almost exactly parallel to the grain, maximising the load carried elastically by the microfibrils and reducing time-dependent deformation during data collection to <10% of the total (Supplementary Figure 1).

The complex group of O-H stretching bands in the FTIR spectra from wood includes contributions from crystalline and disordered cellulose chains, non-cellulosic polysaccharides and water. Initially these interfering contributions were removed by vapour-phase deuteration (Supplementary Figure 2), which substitutes hydroxyl groups on non-cellulosic polymers and on some of the surface cellulose chains[15,32]. The isotope mass effect moves the O-D stretching bands to lower frequency by a factor 1.343. We also used polarised infrared radiation because hydroxyl groups parallel to the fibre axis give longitudinally polarised signals.

The most intense O-H stretching band (or group of bands) remaining after deuteration, the longitudinally polarised band at 3350 $cm^{-1}$, shifted to higher frequency (Supplementary Figure 2) as observed under oscillating stress[32]. The shoulder at about 3280 $cm^{-1}$, also longitudinally polarised, did not shift.

To simplify the spectra further we reversed the deuteration experiment, partially deuterating the crystalline interior of the microfibrils under mild alkaline conditions. Accessible hydroxyls then back-exchanged leaving the internal deuteration stable (Supplementary Figure 3). The Iβ form of crystalline cellulose with hydrogen-bond network A[16] predominated in the deuterated fraction, as shown by a strong shoulder at around 2440 $cm^{-1}$, the absence of a shoulder at 2420 $cm^{-1}$ and low intensity[14,17] above 2510 $cm^{-1}$. Like the O-H stretching band at 3350 $cm^{-1}$, the corresponding intense, longitudinally polarised O-D stretching band at 2490 $cm^{-1}$ shifted to higher frequency under tension and the 2440 $cm^{-1}$ shoulder did not (Figure 2; Supplementary Figure 3).

To take into account some variability between samples in the proportion of the macroscopic strain transferred to cellulose, the longitudinally polarised O-D stretching bandshifts were ratioed against the 1162 $cm^{-1}$ bandshift, a good indicator of the strain on the cellulose microfibrils[32] (Figure 2). Using the difference integral method to quantify bandshifts (Figure 2), a sharp cutoff was evident at 2460 $cm^{-1}$ between shifting and non-shifting bands. No measurable O-D stretching bandshifts were observed in the transversely polarised spectra (Figure 2).

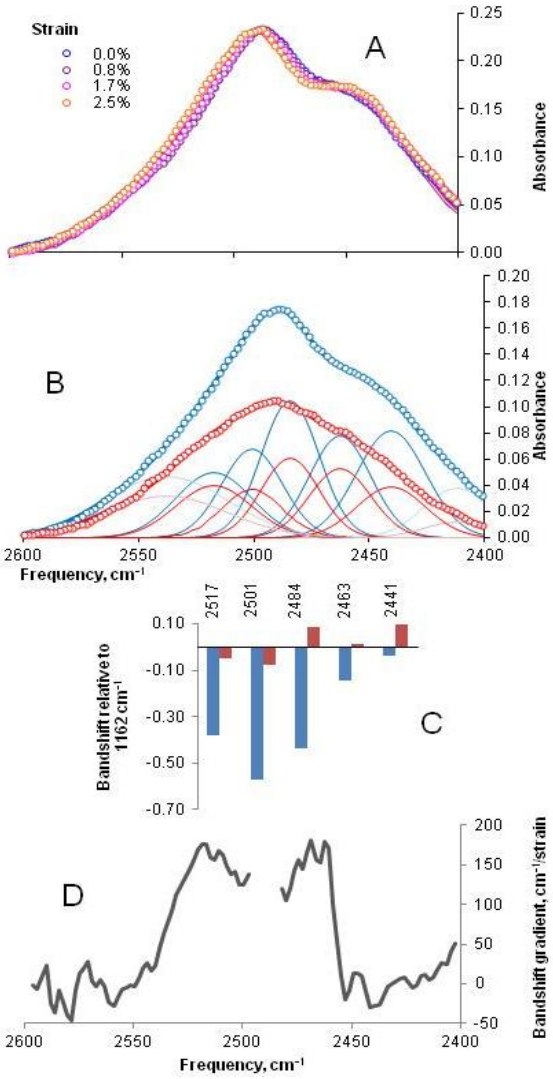



**Figure 2**. (A) Effect of tensile strain on the longitudinally polarised O-D stretching vibrations of internally deuterated spruce wood cellulose. The O-D stretching regions of the longitudinally polarised (blue) and transversely polarised (red) spectra were deconvoluted (B) into five Gaussian bands (with two additional bands fitted to the high- and low-frequency tails). Individual strain–induced bandshifts (C) for the deconvoluted bands (mean of five experiments). Bandshifts for all longitudinally polarised bands were significant ($P<0.05$) except for 2441 $cm^{-1}$ and 2463 $cm^{-1}$. Bandshifts for all transversely polarised bands were non-significant. (D) Spectral variation in bandshift quantified by the difference integral method. A sharp change in local bandshift occurred between 2460 $cm^{-1}$ and 2450 $cm^{-1}$.

Thus it appeared that different hydrogen bonds oriented along the line of tension stretched to different extents. Identifying individual hydrogen-bonded hydroxyls in the FTIR spectra of cellulose is not simple.. By combining polarisation, internal deuteration and matching against DFT predictions[15] we concluded that the shifting bands above 2490 $cm^{-1}$ (Figure 2) corresponded to vibrational modes dominated by O3-D and O6-D. The bands that did not shift, below 2460 $cm^{-1}$, corresponded to vibrational modes dominated by O2-D and O6-D. Since O6-D is oriented transverse to the fibre axis in the predominant hydrogen-bond network A[16], the longitudinally polarised spectra had a reduced contribution from the O6-D component. Thus it was concluded that O3-D was the principal contributor to the deconvoluted bands at 2501 $cm^{-1}$ and 2517 $cm^{-1}$, both of which shifted strongly in the longitudinally polarised spectra (Figure 2), and O2-D was the principal contributor to the 2441 $cm^{-1}$ band which did not shift significantly. These data demonstrate that the O3'H··O5 hydrogen bond lengthened under strain, while the O2H··O6' hydrogen bond did not lengthen. There are implications for the geometry of the stretching chain.

In principle tension can elongate a cellulose chain, not only by stretching the glucose rings and the glycosidic linkages, but also by pulling straight the zig-zag chain conformation in the ring plane. This stretching geometry is consistent with predictions from modelling studies[9,11,12]. If each glucose unit pivots on its linkage oxygens O1 and O4 (Figure 3) the O3'H··O5 hydrogen bond should lengthen

and the O2H··O6’ hydrogen bond should shorten[9].

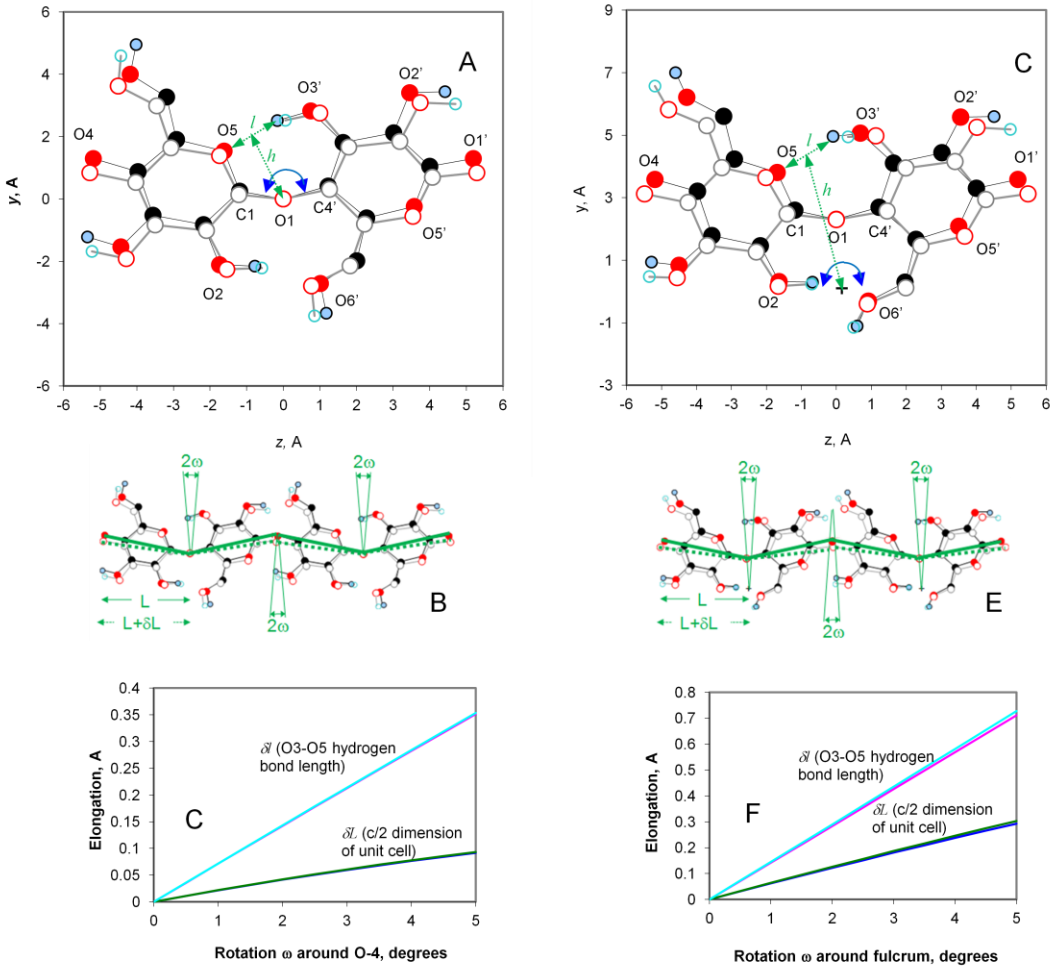



**Figure 3.** Straightening out the kink in the disaccharide unit of cellulose Iβ (origin chain, hydrogen bond network A[16]). The unstrained form is shown (filled circles) in (*a*,*b*) projection in panels A and D. Under tension the two glucose units (open circles) are rotated in opposite directions through an angle ω around the fulcrum of the glycosidic oxygen (O1, centre) (panel A), stretching the O3'H-O5 hydrogen bond by $\delta l = 2h\tan(\omega)$ where $h$ is the projected distance from the fulcrum to the midpoint of the hydrogen bond. Panel B shows how the chain length per glucose unit, $L$ (= c/2), then increases by $\delta L = L(\cos(10.2°-\omega)-\cos\omega)$ as the initial kink of (2 x 10.2°) in the chain is reduced by 2ω. In panel (D) the fulcrum is moved to the midpoint of the O2H-O6' hydrogen bond, since this hydrogen bond was not observed to change significantly in length. With this geometry the glycosidic linkage is stretched at the same time as the O3'H-O5 hydrogen bond, i.e. the C1 to C4 distance increases at the same time as the O3-O5 distance. With this geometry it is simpler to calculate $\delta l$ and $\delta L$ numerically from the atomic coordinates in *(a,b)* projection, and the result is shown for both geometries in Panels C and F. The calculation was carried out for both origin and centre chains in the cellulose Iβ structure but the resulting elongation curves for the two chain types were almost superimposed. The leverage ratio is equal to $\delta l$ / $\delta L$.

We found experimental evidence for lengthening of the O3'D··O5 hydrogen bond but not for significant contraction of the O2D··O6' hydrogen bond. This

implies that the chain does straighten, but simultaneously the glycosidic linkage itself stretches, cancelling out the compression of the O2D··O6' hydrogen bond. An alternative explanation might be that the pivot point remains at the linkage oxygen but that rotation around the C5-C6 bond allows the length of the O2D··O6' hydrogen bond to remain constant. However rotation of C6 would lead to a change in the polarisation of the symmetric and antisymmetric C6-$H_2$ stretching vibrations[28] at 2840-2850 $cm^{-1}$ and 2930-2970 $cm^{-1}$. No such changes in polarisation were observed (Supplementary Figure 4), suggesting that the effective position of the pivot point is probably not at the linkage oxygen but in the region of the O2D··O6' hydrogen bond (Figure 3D).

The two-dimensional depiction in Figure 3 is of course simplified: the changes in glycosidic geometry are likely to be more complex than can be accounted for by a simple, unique pivot point[9,12] since for the projected zig-zag angle of the chain to decrease, the glycosidic torsion angles and the C-O-C bond angle must change and these are not coplanar with the figure. Stretching of the monosaccharide rings is also possible, but is predicted[9] to be considerably less than the stretching of the glycosidic linkage between them.

Due to its position on the flank of the glycosidic linkage, the O3'H··O5 hydrogen bond is well placed to resist the straightening and consequent elongation of the chain. We can speak of this effect as molecular 'leverage' for the hydrogen bond (cf. 'atomic levers'[33]) meaning co-operative action with a fulcrum provided by more rigid covalent bonding. The simplest general way to calculate the effective leverage is from the ratio of the elongation of the O3'H··O5 hydrogen bond $\delta l$ to the chain elongation per monomer unit $\delta L$. These elongations are shown in Figure 3. The leverage $\delta l$ / $\delta L$ varies according to the inverse cosine of the angle of rotation. If the fulcrum is at the glycosidic oxygen the leverage is approximately 4.4. If the fulcrum is at the centre of the O2H··O6' hydrogen bond the leverage falls to about 2.4. This means that the O3'H··O5 hydrogen bond stretches by 2.4 times as much as the monomer length along the chain, and is about 2.4 times as effective in resisting the stretching of the chain as it would be if it were simply stretching in parallel with the covalent linkage and the contributions of the covalent linkage and the hydrogen bond were additive. The leverage is decreased by the slight extensibility of the covalent linkage.

Modelling studies indicate that the O2H··O6' hydrogen bond too is required for increased chain stiffness[9]. Since it does not change significantly in length, it may resist twisting of the chain out of the flat conformation that is optimal for O3'H··O5 hydrogen bonding on the other side of the glycosidic link[9,11], or there may be stereoelectronic synergism along the line of alternating O3'H··O5 and O2H··O6' hydrogen bonds on the same side of the chain.

This leverage mechanism for the stretching of cellulose leads to the testable prediction that the overall width of the zig-zag cellulose molecule will be reduced on elongation. Because the transversely polarised FTIR spectra gave no evidence that transverse hydrogen bonds between chains changed in length (Figure 2), any change in the overall width of the chains under tension should lead to a reduction in the *b* dimension of the unit cell (the dimension across a sheet of chains, assuming the cellulose Iβ lattice). Transverse contraction of the unit cell in the *a* dimension (inter-sheet spacing) under axial stress has previously been observed[34], but changes were not reported in the *b* dimension.

To test that prediction, X-ray diffraction experiments were performed under tension (Figure 4), measuring all three unit cell dimensions. There was more stress

relaxation than in the FTIR experiments due to the longer measurement duration, but approximately half of the applied macroscopic strain was recovered as crystallographic strain, i.e. as an increase in the axial dimension *c* of the unit cell measured from the 004 reflection. This reflection did not broaden significantly under tension (Figure 4), showing that the load distribution between microfibrils remained relatively uniform.

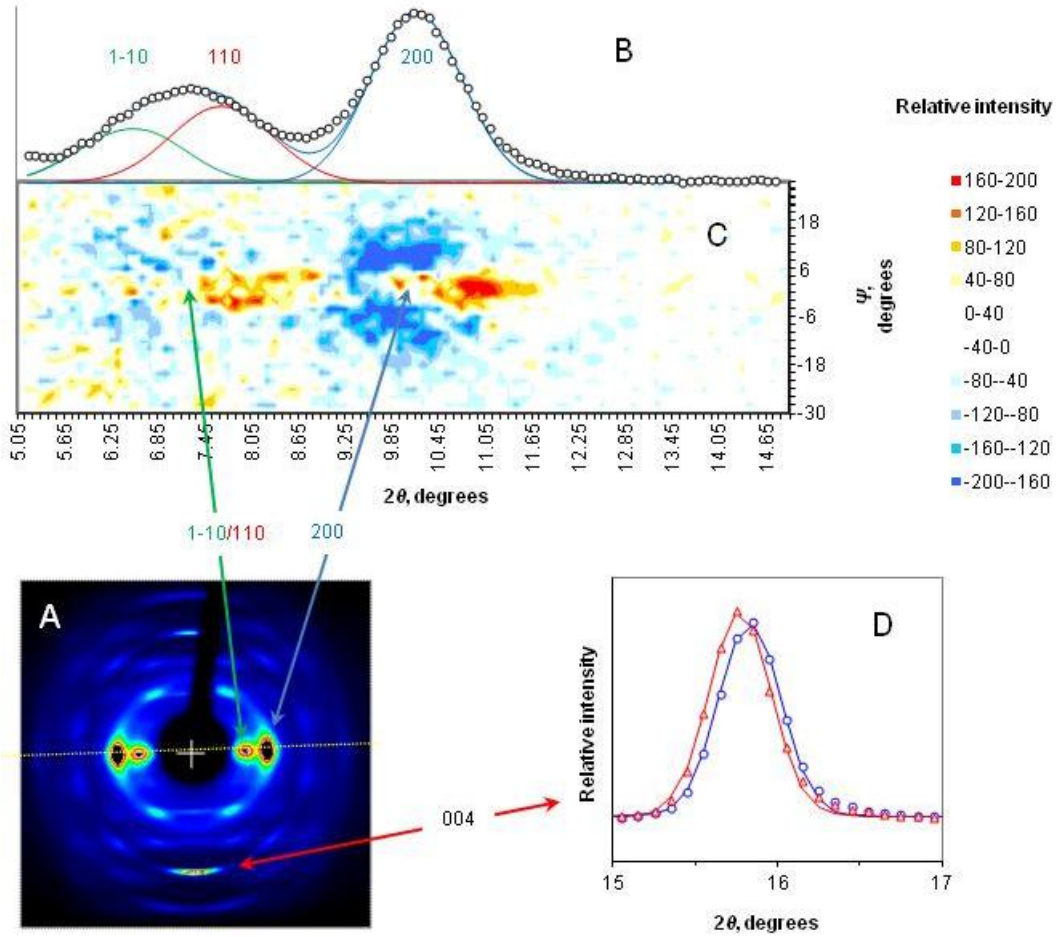



**Figure 4.** Tension-induced changes in the X-ray diffraction pattern (A) from spruce wood cellulose. (B) Equatorial intensity profile, measured along the white dotted line in panel A, with the 1-10, 110 and 200 reflections indexed on the cellulose Iβ lattice. (C) 2D plot of strain-induced intensity changes along the radial profile. Scattered intensity moved from the wings inward towards the centre of the equatorial profile ($\chi$=0) at the same time as the 200 and 110 reflections moved to greater $2\theta$. (D) Changing position of the axial 004 reflection under tension, without change in width: blue, zero strain; red, 1% strain.

The associated changes in lateral dimensions could then be deduced from the equatorial reflections, although the deduction was complicated by increased uniformity of orientation under strain and by the strong overlap between the 1-10 and 110 reflections (Figure 4). The contraction previously observed[34] in the *a* dimension

was evident from the outward displacement of the 200 reflection. The overlapped 1-10 and 110 reflections were also displaced outward and their separation increased[34] . This implies contraction of the *b* dimension across the sheets of chains and an increase in the monoclinic angle as the unit cell elongated (Figure 5). The contraction in the *b* dimension corroborates the mechanism proposed above.

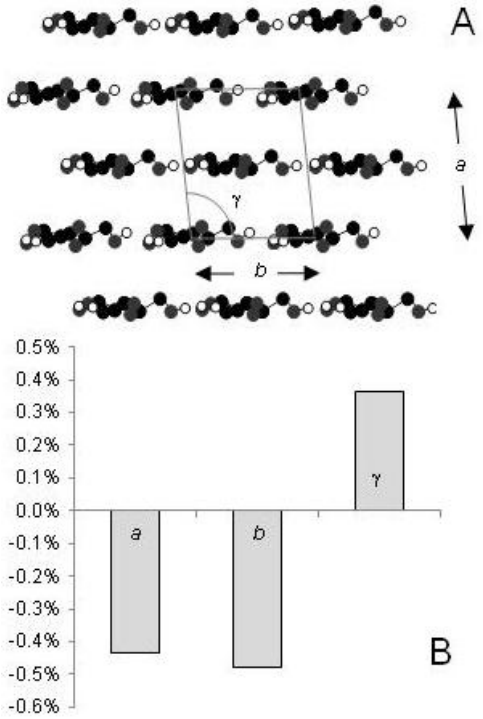



**Figure 5.** The diagram (A) shows a microfibril cross-section with the unit cell outlined. The *c* dimension is perpendicular to the plane of the figure. (B) Strain-induced percentage contraction in the transverse dimensions *a* and *b* of the cellulose Iβ unit cell and increase in the monoclinic angle γ, for a 1% elongation in the axial dimension *c* of the unit cell. The percentage elongation of the *c* dimension (crystallographic strain) was calculated from the change in position of the 004 reflection on the fibre axis. Percentage change in dimension *a* was calculated from the change in position of the 200 reflection, and percentage changes in γ and *b* were then calculated from the positions of the 1-10 and 110 reflections. Pooled data from 6 experiments: percentage changes were significant at P=0.000 (*a* vs *c*); P=0.04 (*b* vs *c*); P=0.004 (γ vs *c*).

## DISCUSSION

The FTIR and diffraction data support a mechanism for elastic extension of the cellulose chain where much of the additional chain length is obtained by straightening out the kink in the chain at each glycosidic linkage[9,12]. The FTIR bandshifts show how this straightening of the chain is resisted by the O3'H··O5 hydrogen bond. Just as a rope can be tightened with great force by pulling sideways on its centre, this geometry provides leverage for the hydrogen bond in restraining the extension of the chain. The geometry is established by the covalent structure of the glycosidic linkage, but not quite rigidly: the lack of any measureable O2-H stretching bandshift shows that the covalent linkage stretches slightly, reducing the leverage for the O3'H··O5 hydrogen bond a little. Co-operation of this kind between covalent and hydrogen bonding, to increase the tensile stiffness of a molecule, has not to our knowledge been previously described. It does not lead to the breaking of the hydrogen bond, until the glycosidic linkage itself breaks.

In these respects cellulose contrasts with spider silk and related strong proteins where sacrificial hydrogen bonds permit controllable stiffness to be combined with high fracture energy[35]. Molecular leverage has the additional effect of stretching the O3'H-O5 hydrogen bond through a much greater amplitude than the cellulose chain itself stretches, and transverse atomic displacements are also large: the resulting dipolar changes may drive the piezoelectric properties of cellulose, which have been exploited in 'smart' cellulose-based devices[36] and have unexplored potential as a mechanism for electromechanical signalling in plants.

There is an unexpected qualitative parallel between the crystallographic effects of strain (Figure 4) and the effects of hydration in wood[31]. Hydration disrupts intramolecular hydrogen bonding in accessible cellulose chains[37]. It also affects stress transmission between fibres, and both mechanisms may contribute to the reduced stiffness of wet wood[31,38], paper and cotton. The rigidity of cellulose chains in different solvents influences the insolubility of microfibrils and their recalcitrance during biofuel production[39].

Isolated cellulose microfibrils have promise for high-performance sustainable nanocomposites[40]. To predict the engineering properties of such materials, continuum mechanical modelling needs to be interfaced with molecular-scale modelling: our analysis shows that a classical mechanics approach at the molecular scale can be surprisingly useful if the system is considered as a nanostructure rather than a continuous material, and gives direct insight into the molecular origins of Poisson ratios.

The mechanism of tensile deformation described here applies to other polysaccharides sharing the key structural motif of a (1,4')β-glycosidic linkage flanked by a O3'H-O5 hydrogen bond. In chitin, the most abundant strong polymer in the animal kingdom, the chain conformation required for the O3'H-O5 hydrogen bond is rigidly constrained by intermolecular hydrogen bonding[41]. The plant hemicelluloses vary in stiffness, due to partial acetylation on O-3 and because they lack the O2H-O6' hydrogen bond that flanks the (1,4') β-glycosidic link on the other side[42].

The concept of synergism between covalent and hydrogen bonding through molecular leverage is in principle applicable also to other proteins, nucleic acids and certain synthetic polymers, but in a less simple form. In these polymers the intramolecular hydrogen-bond donor and acceptor atoms are separated by larger numbers of covalent bonds so that the intervening chain segment is less rigid, a single fulcrum atom cannot normally be identified, and larger clusters of hydrogen bonds are likely to act together[43,44]. The experimental approach described here, using vibrational bandshifts to follow the stretching of individual hydrogen bonds, should be applicable to other polymers.

## CONCLUSION

We conclude that the stiffness of cellulose is enhanced by hydrogen bonding between O3 of one glucose unit and O5 of the preceding glucose unit. The degree of enhancement is substantially greater than the hydrogen bond in question would provide without the molecular leverage effect provided by the geometry of the covalent linkage between the two glucose units. That is, the covalent and hydrogen-bonding systems work in synergy to enhance the mechanical properties of the cellulose chain.

## ASSOCIATED CONTENT

Supporting information: Figures S1-S4. This material is available free of charge via the Internet at http://pubs.acs.org

## AUTHOR INFORMATION

Corresponding author

Email: Michael.Jarvis@Glasgow.ac.uk.

## ACKNOWLEDGEMENTS

We thank C. Lee and J. Kubicki (Pennsylvania State University) for discussions and access to unpublished data. We also thank Y. Nishiyama (CERMAV, Grenoble) for critical comments. This work was supported by the Engineering and Physical Sciences Research Council (UK).

REFERENCES

(1) Fernandes, A. N.; Thomas, L. H.; Altaner, C. M.; Callow, P.; Forsyth, V. T.; Apperley, D. C.; Kennedy, C. J.; Jarvis, M. C. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, E1195.

(2) Jarvis, M. *Nature* **2003**, *426*, 611.

(3) Nishiyama, Y. *J. Wood Sci* **2009**, *55*, 241.

(4) Diddens, I.; Murphy, B.; Krisch, M.; Mueller, M. *Macromolecules* **2008**, *41*, 9755.

(5) Nishino, T.; Takano, K.; Nakamae, K. *J. Polym. Sci., Part B: Polym. Phys.* **1995**, *33*, 1647.

(6) Saito, T.; Kuramae, R.; Wohlert, J.; Berglund, L. A.; Isogai, A. *Biomacromolecules* **2013**, *14*, 248.

(7) Keckes, J.; Burgert, I.; Fruhmann, K.; Muller, M.; Kolln, K.; Hamilton, M.; Burghammer, M.; Roth, S. V.; Stanzl-Tschegg, S.; Fratzl, P. *Nat. Mater.* **2003**, *2*, 810.

(8) Montero, C.; Clair, B.; Almeras, T.; van der Lee, A.; Gril, J. *Compos. Sci. Technol.* **2012**, *72*, 175.

(9) Cintron, M. S.; Johnson, G. P.; French, A. D. *Cellulose* **2011**, *18*, 505.

(10) Sturcova, A.; Davies, G. R.; Eichhorn, S. J. *Biomacromolecules* **2005**, *6*, 1055.

(11) Tashiro, K.; Kobayashi, M. *Polymer* **1991**, *32*, 1516.

(12) Wohlert, J.; Bergenstrahle-Wohlert, M.; Berglund, L. A. *Cellulose* **2012**, *19*, 1821.

(13) Salmen, L.; Bergstrom, E. *Cellulose* **2009**, *16*, 975.

(14) Horikawa, Y.; Sugiyama, J. *Biomacromolecules* **2009**, *10*, 2235.

(15) Thomas, L. H.; Forsyth, V. T.; Sturcova, A.; Kennedy, C. J.; May, R. P.; Altaner, C. M.; Apperley, D. C.; Wess, T. J.; Jarvis, M. C. *Plant Physiol.* **2013**, *161*, 465.

(16) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074.

(17) Sturcova, A.; His, I.; Apperley, D. C.; Sugiyama, J.; Jarvis, M. C. *Biomacromolecules* **2004**, *5*, 1333.

(18) Liang, C. Y.; Marchessault, R. H. *J. Polym. Sci.* **1959**, *37*, 385.

(19) Marechal, Y.; Chanzy, H. *J. Mol. Struct.* **2000**, *523*, 183.

(20) Sugiyama, J.; Persson, J.; Chanzy, H. *Macromolecules* **1991**, *24*, 2461.

(21) Fengel, D. *Holzforschung* **1992**, *46*, 283.

(22) Hinterstoisser, B.; Salmen, L. *Cellulose* **1999**, *6*, 251.

(23) Kokot, S.; Czarnik-Matusewicz, B.; Ozaki, Y. *Biopolymers* **2002**, *67*, 456.

(24) Ivanova, N. V. *Zh. Prikl. Spektrosk.* **1989**, *51*, 301.

(25) Gritsan, V. N. Z., R.G.; Kachur, V.T. *Acta Polym.* **1982**, *33*, 20.

(26) Gardner, K. H.; Blackwell.J *Biopolymers* **1974**, *13*, 1975.

(27) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300.

(28) Lee, C. M.; Mohamed, N. M. A.; Watts, H. D.; Kubicki, J. D.; Kim, S. H. *J. Phys. Chem. B* **2013**, *117*, 6681.

(29) Sturcova, A.; Eichhorn, S. J.; Jarvis, M. C. *Biomacromolecules* **2006**, *7*, 2688.

(30) Thomas, L. H.; Altaner, C. M.; Jarvis, M. C. *J. Appl. Crystallogr.* **2013**, *46*, 972.

(31) Zabler, S.; Paris, O.; Burgert, I.; Fratzl, P. *J. Struct. Biol.* **2010**, *171*, 133.

(32) Hofstetter, K.; Hinterstoisser, B.; Salmen, L. *Cellulose* **2006**, *13*, 131.

(33) Marszalek, P. E.; Pang, Y. P.; Li, H. B.; El Yazal, J.; Oberhauser, A. F.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 7894.

(34) Nakamura, K.; Wada, M.; Kuga, S.; Okano, T. *J. Polym. Sci., Part B: Polym. Phys.* **2004**, *42*, 1206.

(35) Keten, S.; Xu, Z.; Ihle, B.; Buehler, M. J. *Nat. Mater.* **2010**, *9*, 359.

(36) Csoka, L.; Hoeger, I. C.; Rojas, O. J.; Peszlen, I.; Pawlak, J. J.; Peralta, P. N. *ACS Macro Lett.* **2012**, *1*, 867.

(37) Bergenstrahle, M.; Wohlert, J.; Himmel, M. E.; Brady, J. W. *Carbohydr. Res.* **2010**, *345*, 2060.

(38) Obataya, E.; Norimoto, M.; Gril, J. *Polymer* **1998**, *39*, 3059.

(39) Rabideau, B. D.; Agarwal, A.; Ismail, A. E. *J. Phys. Chem. B* **2013**, *117*, 3469.

(40) Klemm, D.; Kramer, F.; Moritz, S.; Lindstrom, T.; Ankerfors, M.; Gray, D.; Dorris, A. *Angew. Chem., Int. Ed.* **2011**, *50*, 5438.

(41) Nishiyama, Y.; Noishiki, Y.; Wada, M. *Macromolecules* **2011**, *44*, 950.

(42) Scheller, H. V.; Ulvskov, P.; Merchant, S., Briggs, W. R., Ort, D., Eds. *Hemicelluloses* 2010; Vol. 61, p 263.

(43) Bosaeus, N.; El-Sagheer, A. H.; Brown, T.; Smith, S. B.; Akerman, B.; Bustamante, C.; Norden, B. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 15179.

(44) Livesay, D. R.; Huynh, D. H.; Dallakyan, S.; Jacobs, D. J. *Chem. Cent. J.* **2008**, *2*.